Density Guided MD-Rosetta Protocol for Protein Structure Refinement

Sumudu Leelananda and Steffen Lindert

Department of Chemistry and Biochemistry, The Ohio State University

Abstract

We have developed a combined Molecular dynamics (MD)-Rosetta protocol that incorporates cryo-electron microscopy data to improve protein structure refinement. This method uses two complementary approaches: Rosetta density guided model refinement via iterative local rebuilding of loops, and molecular dynamics flexible fitting methods. This combined MD-Rosetta protocol shows improvement over refinement with the individual methods.

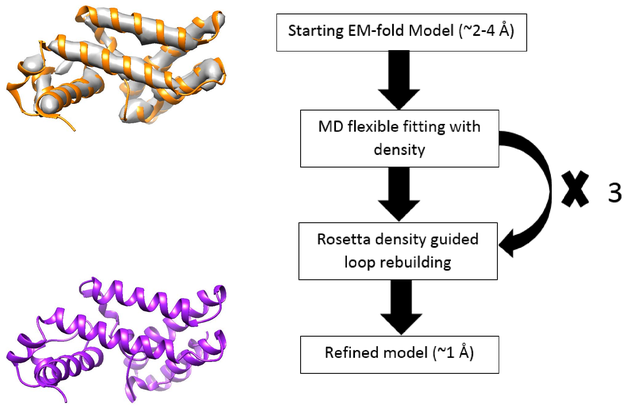

Figure 1: Schematic diagram of the flow of the MD-Rosetta iterative protocol.

Introduction

Success of *ab initio* structure prediction is limited to small proteins. Recently, sparse experimental restraints have been used to predict and refine large protein structures. Incorporation of cryoEM density information has proven to be successful in protein structure prediction and refinement. Even though Rosetta scoring function can successfully distinguish between native-like and non-native-like structures when supported by the density map, one limitation is its inability to sufficiently sample the conformational space to find near native structures. By using molecular dynamics simulations it is possible to overcome this limitation of conformational sampling. The orthogonal sampling and scoring strategy provided by our iterative density-guided MD-Rosetta protocol makes it possible to significantly refine the four benchmark proteins that we tested.

Methods

- Starting structures of four proteins (150-234 residues) with low RMSD (1X91, 1ICX, 1DVO and 2FD5) were generated using EM-Fold.
- Medium resolution density maps were simulated (4-10 Å).
- All RMSDs were calculated using the BCL program.
- First, all the systems were prepared for molecular dynamics in NAMD.
- MDFF run of a 1 ns was followed by a 0.2 ns minimization of the system into the density map.
- After each MDFF simulation, the last frame of the trajectory was used as input into the Rosetta loop building and refinement protocol, and 5000 models were generated.
- The model that gave the lowest Rosetta full atom energy score was obtained to go to the next phase of the iterative protocol.
- After Rosetta3 the best scoring model was picked as the final model of the MD-Rosetta iterative protocol.

Results

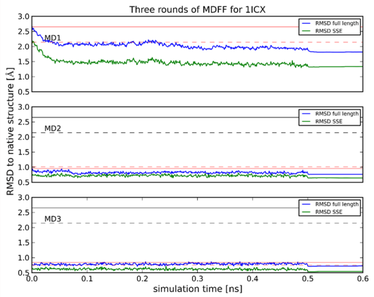

Figure 2: The RMSD of the structure with respect to the native along MD trajectories shown for all 1ICX residues (blue) and for residues in secondary structure elements (green). RMSDs of reference models are displayed by vertical lines: the full length RMSD of the starting model (black line), the RMSD over SSEs of the starting model (dashed black line), the full length RMSD of the best scoring model from the previous Rosetta round (red line) and the RMSD over SSEs of the best scoring model from the previous Rosetta round (dashed red line). For the first round of MD, the red and black lines coincide.

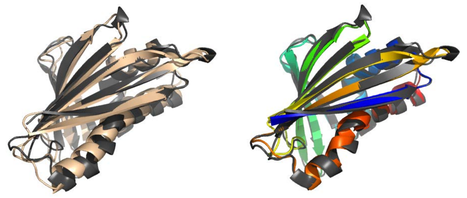

Figure 3 (a) The starting 1ICX model used in the protocol is shown in tan (RMSD to native 2.65 Å). (b) After three MD-Rosetta iterations the final refined model obtained is shown in rainbow-colors (RMSD to native 0.76 Å). The native structure of 1ICX is shown in gray.

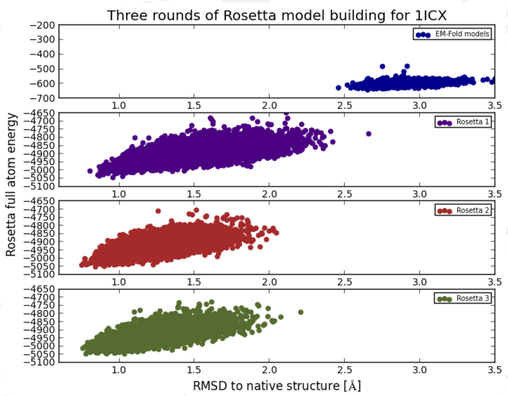

Figure 4: RMSD vs score plots for 1ICX. The first panel (blue) shows the starting EM-Fold models and the other three panels show the results for the first (purple), second (red) and third (green) Rosetta round of the iterative MD-Rosetta.

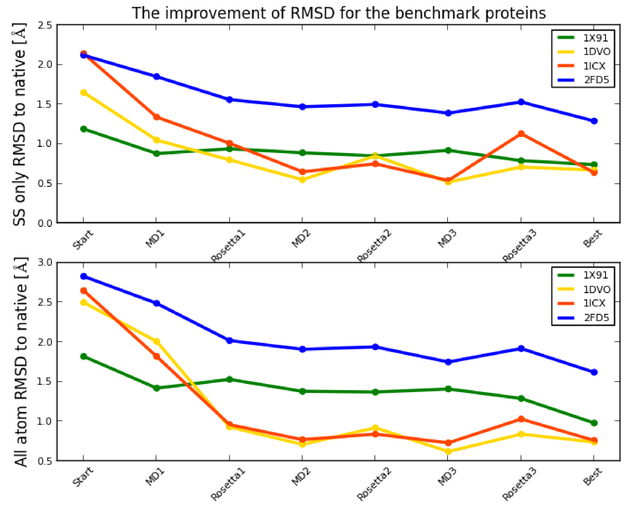

Figure 5: The improvement of the RMSD of the structures for 1X91 (green), 1DVO (yellow), 1ICX (red) and 2FD5 (blue) over the MD and Rosetta steps of the iterative protocol. The first panel shows all heavy atom RMSD with native and the second panel shows RMSD for only the secondary structure elements.

Conclusion

- Medium resolution cryoEM guided iterative refinement protocol significantly improves the model quality of the four benchmark proteins.
- Molecular dynamics simulations help escape conformational traps and better sample the conformational space.
- These simulations are most efficient in refining secondary structure elements of the models.
- Rosetta refinement is more powerful in refining side chains and loop regions.
- The hybrid MD-Rosetta protocol could be potentially used in improving model quality of various structures determined by cryoEM guided experiments.

References

Lindert, S.; Meiler, J.; McCammon, J. A., J Chem Theory Comput 2013, 9 (8), 3843-3847.

Lindert, S.; McCammon, J. A., J Chem Theory Comput 2015, 11 (3), 1337–1346.